**Molecular Analysis of the *Corynebacterium glutamicum* Transketolase Gene**

Masato Ikeda,† Nozomu Kamada, Yutaka Takano, and Tetsuo Nakano

Technical Research Laboratories, Hofu Plant, Kyowa Hakko Kogyo Co., Ltd., 1-1 Kyowa-machi, Hofu, Yamaguchi 747-8522, Japan

Received April 13, 1999; Accepted June 3, 1999

Transketolase is important in production of the aromatic amino acids in *Corynebacterium glutamicum*. The complete nucleotide sequence of the *C. glutamicum* transketolase gene has been identified. The DNA-derived protein sequence is highly similar to the transketolase of *Mycobacterium tuberculosis*, taxonomically related to *C. glutamicum*. The alignment of the N-terminus regions between both transketolases showed TTG to be the most probable start codon. Potential ribosomal binding and promoter regions were situated upstream from the TTG. The deduced amino acid sequence consists of 700 residues with a calculated molecular mass of 75 kDa, and contains all amino acid residues involved in cofactor and substrate binding in the well-characterized yeast transketolase sequence.

**Key words:** *Corynebacterium glutamicum*; transketolase; *tkt* gene

Transketolase (EC 2.2.1.1), ubiquitous in animals, plants, and microorganisms, is a key enzyme of the nonoxidative pentose phosphate pathway. The enzyme requires thiamin pyrophosphate as a cofactor, as well as magnesium ions for its activity. It catalyzes the reversible transfer of a ketol group between a ketose phosphate and an aldose phosphate. Together with aldolase, transketolase creates a reversible link between glycolysis and the pentose phosphate pathway, thereby enabling the cells to shuttle ribose 5-phosphate and glycolytic intermediates between the two pathways.

Microbial transketolases and their genes have been studied in limited sources including *Escherichia coli*, *Rhodobacter sphaeroides*, and yeasts. In *E. coli*, two transketolase isozymes encoded by the *tktA* and *tktB* genes have been reported. The *tktA* gene product, the major isozyme, shows 74% amino acid identity with the *tktB* gene product. Similarly, there are two transketolase isozymes in *Saccharomyces cerevisiae*. Amino acid sequences of these transketolases have shown that they have similar subunits with molecular masses of 70–80 kDa.

Recently, we isolated a transketolase-negative mutant of *Corynebacterium glutamicum* to examine the physiological role of the enzyme in the organism, which showed that a single transketolase was responsible for biosynthesis of the aromatic amino acids. Furthermore, we cloned the transketolase gene of wild-type *C. glutamicum* ATCC 31833 and showed that the overexpressed transketolase could function in directing carbon toward erythrose 4-phosphate formation, leading to increased production of the aromatic amino acids. However, structural characteristics of the transketolase gene remain to be analyzed. Here we describe the molecular characterization of the gene in order to advance the molecular biology of this industrially important organism as well as to gain basic information on the functional domains and on the transcription and translation signals.

The transketolase gene of *C. glutamicum* was originally cloned as a 7.6-kb EcoRI fragment and was later subcloned into a 3.2-kb XhoI-SalI fragment. To localize the transketolase gene more precisely within the 3.2-kb fragment, several subclones were constructed with the use of the high-copy-number vector pCG116 and then examined for complementation ability of *C. glutamicum* RA60, a transketolase mutant with a shikimic acid-requiring phenotype. Transketolase activity in the crude cell extract of each clone was also measured as described previously.

The results of these subcloning experiments indicated that the functional transketolase gene was localized within the 2.4-kb *AffII* fragment (Fig. 1). Deletion of the 0.16-kb *AffII*-*BglII* region of the left-end within the 2.4-kb fragment resulted in a decrease of about 90% in the transketolase activity although this still allowed it to complement the shikimic acid auxotrophy of strain RA60 (data not shown).

**Fig. 1.** Restriction Map and Sequencing Strategy of the 2433 bp *AffII* Fragment Containing the *C. glutamicum* Transketolase Gene.


† To whom correspondence should be addressed. Masato Ikeda, FAX: +81-835-22-2466; E-mail: m.iokeda@kyowa.co.jp
with respect to the vector and was used to analyze the complete DNA sequences of both strands. DNA sequencing was done by the dideoxy chain termination method of Sanger\(^1\) using universal primers. The sequencing strategy and a restriction map are depicted in Fig. 1. The nucleotide sequence and the deduced amino acid sequence of the transketolase gene are shown in Fig. 2. The DNA sequence has been submitted to DDBJ/GenBank/EMBL and has been assigned the accession number AB023377.

Computer-assisted analysis found only one large open reading frame within this fragment. To deduce the translational start codon, the predicted amino acid sequences was aligned with the transketolase sequence of Mycobacterium tuberculosis because Corynebacterium and Mycobacterium genera both belong to the actinomycete

---

Fig. 2. Nucleotide Sequence of the 2433 bp AflII Fragment Containing the C. glutamicum Transketolase Gene and Its Deduced Amino Acid Sequence. The putative promoter region (\(-35\) and \(-10\) regions) and the proposed ribosomal binding site (RBS) are indicated. The putative terminator is shown by inverted arrows.

**Fig. 3.** Alignment of the Amino Acid Sequences of Transketolases from Different Sources by Using the Clustal W (1.7) Program. C. glut., *C. glutamicum* transketolase; M. tube., *M. tuberculosis* *kII* product; E. co-A., *E. coli* *kII* A product; S. ce-1, *S. cerevisiae* transketolase 1; R. sphaeraeides, *R. sphaeraeides* transketolase. Asterisks (*) and dots (•) indicate identical and similar residues, respectively. Signs above the sequence that were identified as essential for cofactor (•) or substrate (○) binding of the yeast enzyme.
phylogenetic branch and are closely related. TTG at position 224 was found as the most probable translational start codon since a polypeptide translated from this TTG shows the best N-terminal similarity to the transketolase from *Mycobacterium tuberculosis* (Fig. 3) and since a possible ribosomal binding site (5'-AGGA-3') was found 4 bp upstream (Fig. 2). An in-frame TTG codon (position 233) lies 6 bp downstream of the proposed start site (Fig. 2). This TTG is also a potential candidate for the translational start site since a second possible ribosomal binding site (5'-AGGA-3') was situated 9 bp upstream. Although TTG is an unusual start codon, a number of bacterial genes have been shown to initiate with this codon, including the argC gene from *Bacillus subtilis*, the endoglucanase gene from *Clostridium acetobutylicum*, and the pyruvate kinase gene from *Zymomonas mobilis*. Another potential translational start site within this gene, GTG at position 296, is a less likely candidate for the translational start site since the codon lies downstream of the region of significant sequence similarity with the *Mycobacterium tuberculosis* transketolase gene product and in addition is not preceded by a well-spaced ribosomal binding site. Assuming translational initiation at TTG at position 224, the *C. glutamicum* transketolase gene product consists of 700 amino acid residues with a calculated molecular mass of 75 kDa. This is in close agreement with the estimated molecular weight of the overproduced transketolase protein after SDS/polyacrylamide gel electrophoresis. The open reading frame is followed by a potential stem-loop structure indicative for a rho-independent terminator, which is located 21 bp downstream of the TAA stop codon at position 2234 (Fig. 2). The free energy of this terminator spanning nucleotides 2348–2393 is ΔG° of −172 kJ/mol.

Analysis of the nucleotide sequence at the 5' end of the transketolase gene found that the potential −35 region (5'−TAGATC−3') and −10 region (5'−GAAAC−3') promoter regions are appropriately situated upstream of the proposed start site (Fig. 2). The putative −35 promoter region is overlapped with a BglII site (Fig. 2), which might reasonably explain the drastic decrease in transketolase activity after deletion of the 0.16 kb *AflII-BglII* region of the left-end within the 2.4-kb *AflII* fragment (described above).

The predicted amino acid sequence of the *C. glutamicum* transketolase gene product was aligned with the transketolase sequences of other origins (Fig. 3), which revealed high similarities with the transketolase from *M. tuberculosis* (63% identical residues), the *tklA* product of *E. coli* (45%), the *tkl* product of *S. cerevisiae* (42%), and the transketolase from *Rhodobacter sphaeroides* (43%).

These alignment studies demonstrated several regions of strongly conserved amino acid residues (Fig. 3). Amino acid residues 34–54 and 500–516 are identical to transketolase signature motifs 1 and 2, respectively, which were found by searching protein sequence motifs in Prosite Pattern of GenomeNet (http://www.genome.ad.jp). The region of motif 1 contains a histidine residue that has been assumed to function in proton transfer during catalysis in the yeast transketolase. On the other hand, motif 2 is a part of the active cleft that has been assumed to participate in substrate-binding in the yeast enzyme, and thus may be the substrate-binding domain. A third conserved region, 179–210, is similar to a consensus sequence [GDG-(X<sub>0</sub>)E(X<sub>0</sub>)A(X<sub>0</sub>)N] previously identified by Hawkins in a survey 16 thiamin pyrophosphate-binding enzymes and is the possible cofactor-binding domain. From X-ray crystallography using the purified yeast enzyme, Lindqvist et al. have shown which amino acid residues are involved in cofactor and substrate binding. All these amino acid residues are found in the *C. glutamicum* transketolase sequence (Fig. 3). These results indicate that both cofactor and substrate binding domains are highly conserved in the *C. glutamicum* enzyme.

In summary, we have presented the structure of the *C. glutamicum* transketolase gene and the deduced amino acid sequence. Furthermore, we have proposed the functional domains of the enzyme by a comparative study with other transketolases such as the yeast enzyme well-characterized for structure-function relationships. These findings will aid the molecular construction of better production strains with fine-tuned transketolase activity.

**Acknowledgment**

We are grateful to Mr. Y. Yonetani of Tokyo Research Laboratories, Kyowa Hakko Kogyo, for his helpful discussions.

**References**